DOI:
10.1037/ccp0000310

Document Version
Peer reviewed version

Citation for published version (APA):
Hirsch, C. R., Krahé, C., Whyte, J., Loizou, S., Bridge, L., Norton, S., & Mathews, A. (Accepted/In press). Interpretation training to target repetitive negative thinking in Generalized Anxiety Disorder and Depression. *Journal of Consulting and Clinical Psychology*, *86*(12), 1017-1030. https://doi.org/10.1037/ccp0000310

Running head: Interpretation training for repetitive negative thinking

Interpretation training to target repetitive negative thinking in Generalized Anxiety Disorder and Depression

Colette R. Hirsch[1], Charlotte Krahé[1], Jessica Whyte[1], Sofia Loizou[1], Livia Bridge[1], Sam Norton[1], and Andrew Mathews[2]

[1]King's College London, London, Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, UK.

[2]Department of Psychology, University of California, Davis, California, USA

Correspondence:

*Dr Colette R. Hirsch

King's College London, Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience,

De Crespigny Park,

London SE5 8AF.

Email: colette.hirsch@kcl.ac.uk

Abstract

Objective: Repetitive negative thinking (RNT) e.g., worry in generalized anxiety disorder (GAD) and rumination in depression, is often targeted during psychological treatments. To test the hypothesis that negative interpretation bias contributes to worry and rumination, we assessed the effects of inducing more positive interpretations in reducing RNT.

Method: Volunteers diagnosed with GAD (66) or Depression (65) were randomly allocated to one of two versions of Cognitive Bias Modification (CBM-I), either with or without RNT priming prior to training), or a control condition, each involving 10 internet-delivered sessions. Outcome measures of interpretation bias, a behavioral RNT task and self-reported worry, rumination, anxiety and depression were obtained at baseline, after home-based training and at 1-month follow up (self-report questionnaires only).

Results: CBM-I training, across diagnostic groups, promoted a more positive interpretation bias and led to reductions in worry, rumination, and depressive symptoms, which were maintained at follow up. Anxiety symptoms were reduced only in the GAD group at follow up. There were no differences between CBM-I versions; brief priming of RNT did not influence CBM-I effectiveness. Level of interpretation bias post training partially mediated the effects of CBM-I on follow-up questionnaire scores.

Conclusions: In contrast to some recent failures to demonstrate improvements following internet-delivered CBM, we found that self-reported RNT and negative mood were reduced by CBM-I. This is consistent with a causal role for negative interpretation bias in both worry and rumination, suggesting a useful role for CBM-I within treatments for anxiety and depression.

Key words: Generalized anxiety disorder (GAD); Depression; interpretation bias; cognitive bias modification (CBM); repetitive negative thinking

Public Health Significance Statements

Many people worry about the future, or mull over negative events from the past (rumination). These types of unhelpful repetitive negative thinking can maintain clinical anxiety and depression. This study indicates that simple regular practice in making positive interpretations of emotionally ambiguous information reduces repetitive negative thinking in individuals with clinical anxiety or depression, and also improves mood.

Interpretation Training to Target Repetitive Negative Thinking in Generalized Anxiety Disorder and Depression

Repetitive negative thinking (RNT) occurs in many emotional disorders, with worry and rumination being the two most obvious examples. Uncontrollable worry about multiple future events is central to the diagnosis of Generalized Anxiety Disorder (GAD), while rumination (repeatedly thinking about past or current concerns) is more often reported (along with worry) in Depression. Both these forms of RNT are characterized by their negative content, an over-general abstract style and – in pathological conditions – their apparently uncontrollable and perseverative nature. These overlapping characteristics, as well as their co-occurrence within individuals and across disorders, have led to them being conceptualized as a transdiagnostic process termed repetitive negative thinking (Drost, van der Does, van Hemert, Penninx, & Spinhoven, 2014).

Although similar in many respects, worry and rumination are sometimes described as differing in content, with worry focused on possible future threats and rumination more likely to concern past/ongoing personal failures. Consequently, it remains unclear whether they are underpinned by the same cognitive mechanisms and, furthermore, whether they can be modified using the same methods. This is of some importance, given the assumed role of RNT in maintaining clinical disorders. For example, rumination prolongs depression episodes (Watkins, 2008) and worry maintains anxiety (Hirsch & Mathews, 2012).

Hirsch and Mathews (2012) identified three critical processes thought to underlie pathological worry, namely emotional processing biases favoring negative information, a verbal thinking style, and deficits in attentional control. In the current study, we focus on emotional processing biases and specifically on negative interpretation bias – the tendency to habitually interpret ambiguous information as negative or threatening - and investigate whether this bias plays a similar causal role in both worry and rumination.

Basic Research on the Nature of Interpretation

Early studies of how ambiguous information is resolved during reading revealed that, in early stages of processing, alternate resolutions are activated prior to one interpretation reaching awareness. Thus, after reading a sentence such as "He played the ace of Spades", one is not usually aware of alternative meanings of "spade", yet the decision to reject "dig" as being related to the sentence is initially slowed, although in proficient readers this interference effect dissipates very rapidly (Gernsbacher & Faust, 1991). These experiments imply that alternative but contextually irrelevant resolutions are typically activated, but are then quickly suppressed prior to awareness. Which meaning of ambiguous information becomes dominant depends partly on context (as in the above example), but is also influenced by its prior frequency of use. Thus, the homograph "growth" is likely to prime "plant" more than "tumor" for a gardener, but probably the converse for an oncologist. Eysenck, Mogg, May, Richards, and Mathews (1991) tested the related hypothesis that emotional disorders are similarly associated with resolutions of ambiguity that are congruent with habitual thought content. Individuals with GAD were more likely than non-anxious controls to interpret ambiguous sentences in terms of the threatening rather than their benign meaning. Similarly, Butler and Mathews (1983), Mathews, Richards, and Eysenck (1989), Mogg, Baldwin, Brodrick, and Bradley (2004), and Anderson *et al.* (2012) identified a negative interpretation bias (henceforth interpretation bias) in people suffering from GAD. Negative interpretations are also evident in clinical and sub-clinically depressed individuals (e.g. Nunn, Mathews, & Trower, 1997; Berna, Lang, Goodwin, & Holmes, 2011), particularly in relation to self-referent information (Wisco & Nolen-Hoeksema, 2010). According to a recent meta-analysis, interpretation bias in depression has a medium effect size (Everaert, Podina, & Koster, 2017). Hence, there is evidence of interpretation biases across depression and GAD.

As discussed above, RNT is common across depression and GAD. Hirsch and Mathews (2012) suggest that worry episodes can be triggered by negative interpretations and, once begun, subsequent interpretations direct worry to increasingly negative content. Suarez and Bell-Dolan (2001) demonstrated that children with higher trait worry generated more negative interpretations. In adults, Mor, Hertel, Ngo, Shachar, and Redak (2014) found that greater levels of rumination were associated with more negative interpretations. In a study related to the current paper (Krahé, Whyte, Bridge, Loizou & Hirsch, submitted), levels of trait worry and rumination were both associated with interpretation bias across individuals with depression or GAD, and community controls, even when controlling for levels of depression and anxiety. Hence, both worry and rumination appear to be related to degree of interpretation bias.

Other research has investigated whether biases of interpretation can be acquired in unselected volunteers by repeated presentation of emotional ambiguity which is then consistently resolved in either a positive or negative direction (e.g., Grey & Mathews, 2000; Mathews & Mackintosh, 2000; see Hertel & Mathews, 2011, and Hirsch, Meeten, Krahé, & Reeder, 2016, for reviews). It has been shown that single-session positive training procedures can result in positive emotional changes (e.g., Hoppitt, Mathews, Yiend, & Mackintosh, 2010). Conversely, consistently reinforcing *negative* interpretations in unselected samples increases state rumination (Hertel, Mor, Ferrari, Hunt, & Agrawal, 2014).

From Experimental Research to Clinical Application

Further research has explored these modification methods in sub-clinical populations. Hindash and Rottenberg (2017) used single-session training to facilitate more benign interpretations in dysphoric individuals and found that this led to reduced stress reactivity. Other single-session studies designed to test the effectiveness of such training in individuals with elevated anxiety or depression have also given positive results. For example, in studies

of participants with high levels of worry or GAD, those allocated to positive training condition not only resolved new descriptions in a more positive manner than those allocated to a control condition, but also reported fewer negative thought intrusions in a subsequent test of worry (Hirsch, Hayes, & Mathews, 2009; Hayes, Hirsch, Krebs, & Mathews, 2010).

Together these findings point to a causal role of interpretation bias in maintaining both worry and rumination. Thus, interpretation bias seems a promising candidate target for interventions designed to reduce both worry and rumination. Although important for our theoretical understanding, single-session CBM experiments do not provide evidence of any sustained impact of changing interpretation bias. Multi-session training over several days or weeks, where bias change is assessed after training, and including a post-training follow-up period, is necessary to investigate the longer-term effectiveness of CBM-I. If successful, this could allow widespread dissemination of these methods via the internet, so potentially reaching many people suffering from anxiety or depression who are unable or unwilling to attend clinics for treatment.

However, whilst a number of multi-session studies CBM studies focusing on interpretation bias have shown promising results in terms of reductions in key symptomatology (e.g., Amir & Taylor, 2012; Lang, Blackwell, Harmer, Davison, & Holmes, 2012, Pictet, Jermann, & Ceschi, 2016; Torkan *et al.*, 2014), some other recent trials of CBM designed to modify processing biases via the internet have produced disappointing findings. Studies of attentional retraining for social anxiety (e.g. Carlbring *et al.*, 2012) and of interpretation bias training in depression (e.g., Blackwell *et al.*, 2015) have resulted in the supposedly active training methods having clinical outcomes no better than alternative control conditions. One key question to be resolved before further trials are conducted (particularly via the internet) concerns the factors needed for more effective and robust training methods. One possible explanation put forward for the failure of previous trials was that emotional

concerns naturally aroused in the clinic are absent during training conducted at home. It remains unclear, however, why this would influence training effects.

Recent animal and human research on memory reconsolidation (e.g., Nader & Hardt, 2009) has shown that changing emotional memories depends critically on their re-activation prior to modification via new learning. In the same way, in the absence of the activation of emotional concerns, habitual emotional biases may be less easy to modify by replacing them with more positive learning experiences. Relating to this explanation, the same training method used in one unsuccessful trial of social anxiety (Carlbring *et al.*, 2012) was repeated but with added instructions being given to engage in a socially challenging task prior to each practice session, and this addition significantly improved the outcome of training (Kuckertz *et al.*, 2014). Although this improvement could be attributed to the additional exposure involved, other studies have suggested that activating concerns via instructions can also serve to enhance subsequent negative biases, consistent with the idea that activation effects may be achieved without actual behavioral exposure (Hertel & El-Messidi, 2006; Williams, Mathews, & Hirsch, 2014). Of course, these preliminary findings are hardly conclusive, and it remains possible that activation of emotional concerns could actually interfere with positive retraining, due to depletion of cognitive resources (Hayes, Hirsch, & Mathews, 2008; Stefanopoulou, Hirsch, Hayes, Adlam, & Coker, 2014) that may be necessary for relearning during CBM. In the present study, we investigated the effects of emotional activation prior to training trials compared with the same training given without such activation, to test whether prior activation of emotional concerns (thinking about worry or rumination-related topics) serves to enhance effects of bias modification (as expected from reconsolidation theory), or has no such effect, or even interferes adversely with positive relearning.

The main aim of the present study was to determine whether interpretation bias contributes to worry in GAD and rumination in Depression. To do this, we evaluated the

effects of interpretation bias training on RNT in groups with either diagnoses of GAD or Depression, to investigate whether such training is similarly effective in reducing both worry and rumination, and also results in corresponding improvements in mood. The training task (based on Mathews & Mackintosh, 2000; Holmes & Mathews, 2005) involves repeated practice in listening to ambiguous event descriptions, which are resolved in a benign manner, each followed by a 'comprehension' question that requires confirmation of the positive resolution. In this way, participants are repeatedly but unobtrusively guided towards anticipating and generating positive resolutions of ambiguous situations. In many previous studies of this type, effects of training have been compared with those of a control condition in which ambiguous descriptions were resolved in a positive direction half of the time and negatively for the other half. However, as discussed in detail by Blackwell, Woud, and MacLeod (2017), selection of an appropriate control condition when attempting to determine the causal role of cognitive processes in psychopathology requires one that does not modify bias. While some multi-session studies using a control condition with 50:50 contingency have demonstrated greater training effects in the active condition than the control (e.g., Pictet *et al.*, 2016), others have failed to do so (e.g., Blackwell et al., 2015). It is possible that this contingency may inadvertently promote change by drawing attention to the possibility of different outcomes, particularly when many training sessions are completed over time, as acknowledged by Blackwell et al. (2015). Hence, we elected to build on Murphy, Hirsch, Mathews, Smith, and Clark (2007)'s control condition in which ambiguity remained unresolved. Consequently, training effects in the present experiment were compared to changes occurring in a group of participants randomly allocated to a control condition involving exposure to the same ambiguous material, but without being guided to either a negative or positive resolution.

The secondary aim of the study was to investigate the modulatory role of engaging in RNT prior to training, by including two positive training groups, one with and one without activation of emotional concerns (via worry or rumination) prior to training sessions, allowing the assessment of any differential effects due to such activation. Together, these design features were intended to answer questions relevant to possible treatment applications of cognitive bias modification as a way of reducing worry and rumination that could be readily accessed via the internet.

Hypotheses

1) First, for both GAD and depression groups, we predicted that CBM-I – with or without prior RNT activation – would promote a more positive interpretation bias and reduce consequent levels of worry and rumination, and psychological distress (levels of anxiety and depression) relative to the active control condition. In addition, we conducted planned subgroup analyses to investigate whether any effects of CBM-I on measures of worry, rumination, anxiety, and depression were diagnosis specific.

2) Second, we expected prior RNT activation to modulate these effects, but examined the direction of this effect in an exploratory manner. In particular, we investigated whether RNT activation prior to CBM-I either enhanced the effects of training (by activating underlying cognitive biases, as in reconsolidation research), or reduced training effects (perhaps due to the additional demands placed on attentional control resources by RNT).

3) Last, given its proposed underlying role, we expected that effects of CBM-I on worry and rumination, anxiety and depression at 1-month follow up would be mediated in part by post-intervention level of negative interpretation bias.

Method

Design

Community volunteers with GAD or depression were randomly allocated to one of three conditions: CBM-I with prior RNT (henceforth CBM_RNT), CBM-I without prior RNT (henceforth CBM_STAND), or an active control condition (henceforth CONTROL). Tasks assessing interpretation bias and a behavioral measure of RNT were administered during study visits prior to and following 10 computerized 'training' sessions. Questionnaire measures of RNT and mood were completed online prior to and following the block of training sessions, and additionally at 1-month follow up (see Krahé, Mathews, Whyte, & Hirsch, 2016 for the experimental protocol for this study, which has been previously published, with no changes to methods or procedures). It should be noted that this study was designed as an experiment to examine the role of interpretation bias in maintaining RNT, rather than as a clinical trial; hence, there was no clinical trial registration for this experiment. A flow chart of the experimental procedure is presented in Figure 1.

[INSERT FIGURE 1 HERE]

Participants

One hundred and fifty-seven participants with GAD or MDD were recruited from the community in Greater London via advertisements on websites and in newspapers, as well as via university circular emails and completed at least the first study visit at King's College London. The CONSORT diagram is presented in Integral Supplementary Materials. Participants had to be fluent in English, with normal or corrected hearing, and between 18 and 65 years old. They were initially screened for levels of anxiety and/or depression, that is, they had to have a total score ≥10 or five items scored ≥ 2 including items 1 and/or 2 on the PHQ-9 (Kroenke & Spitzer, 2002), and/or a total score ≥ 10 and item 2 scored ≥ 2 on GAD-7 (Spitzer, Kroenke, Williams, & Löwe, 2006). Individuals taking psychotropic medication had to be stabilized on that medication for at least 3 months without remission. Exclusion criteria were severe depression (≥ 23 PHQ-9 total score), past or current risk to self (self-harm in past

12 months / suicide attempt in last 5 years / PHQ-9 suicidal ideation item 9 scored > 1;Williams, Blackwell, Holmes, & Andrews, 2013), co-morbid psychosis, bipolar disorder, borderline personality disorder or substance abuse, non-normal / not corrected to normal hearing (as the study involves listening to audio clips), as well as current or recent (past 6 months) psychological treatment. Diagnosis of GAD or MDD was assessed using the Structured Clinical Interview for DSM-V axis I disorders (SCID; First, Williams, Karg, & Spitzer, 2015) during a screening telephone assessment prior to the first study visit. An independent rater coded 20% of SCID assessments to check diagnosis; inter-rater agreement was excellent (Cohen's kappa = .96). Participants with current co-morbid GAD and MDD were excluded. Of the 157 participants who completed visit 1 (baseline visit), 10 subsequently dropped out (6%), 5 completed 7 / 10 online sessions or returned for the second (post-training) visit more than one month after the first visit (and thus did not count as 'completers'; see below), 4 started treatment while enrolled in the study, and 7 were excluded for other reasons. The final sample used in the analysis comprised 131 individuals, with 44 in the CONTROL (22 GAD; 22 Depression) and CBM_STAND (22 GAD; 22 Depression) conditions, and 43 in the CBM_RNT condition (22 GAD; 21 Depression).[1] [2] Participant demographic characteristics across groups are presented in Table 1 (see Additional Supplementary Table 1 for characteristics for each group separately).

[INSERT TABLE 1 HERE]

Experimental Conditions

All conditions involved 10 sessions: one initial visit followed by nine sessions completed at home using a purpose-built online platform over the next three weeks to one

[1]The final sample of 131) did not differ from the 26 excluded participants on any questionnaire measures at visit 1 (t (155) = 0.58, p = .560 for PHQ-9; t (155) = 1.24, p = .218 for GAD-7; t (155) = 0.97, p = .336 for PSWQ; t (155) = 0.84, p = .401 for RRS).

[2] See Krahé *et al*. (2016) for sample size and randomisation process.

month[3]. All online sessions began with either an RNT induction (CBM_RNT) or neutral task (CBM_STAND or CONTROL). Then, participants listened to 50 audio clips (henceforth scenarios), imagined themselves in each described situation, and then answered a comprehension question.

Pre-scenario task: RNT induction or neutral task.

RNT induction. The RNT induction was adapted from Hertel *et al.* (2014). Participants selected one of three themes (see Krahé *et al*., 2016 for details) about which they had found themselves worrying (GAD) or ruminating (MDD) recently. Each theme could be selected only once a week (three times in total) to ensure that all participants engaged in RNT across a variety of themes. Participants wrote a one-line summary of their RNT on the chosen topic and this was displayed on subsequent screens. They then wrote down their usual negative thoughts about the topic for three minutes (akin to Cohen, Mor, & Henik, 2015; Grisham, Flower, Williams, & Moulds, 2011). Finally, they worried/ruminated silently about their topic for two minutes and, as a manipulation check, rated their current level of worry, rumination, anxiety, and depression on 0 to 100% visual analogue scales.

Neutral task. To control for the time taken during the RNT induction, participants in the CBM_STAND and CONTROL conditions completed a neutral task. They read neutral stories and made grammatical correctness judgments. At the end of the stories they also completed comprehension questions and rated their worry, rumination, anxiety, and depression.

Main online scenario-based task.

CBM-I. The CBM_RNT and CBM_STAND conditions required participants to listen to scenarios describing situations relating to common worry-related (GAD group) and

[3] We allowed some leeway to complete any outstanding sessions; thus, the maximum time allowed to complete the sessions was one month (see also below).

rumination-related (MDD group) themes, which were emotionally ambiguous but eventually resolved in either a positive (76% of the time), or negative manner (12%), or were left unresolved (12% test trials) by the final words of the scenario. After each scenario, participants completed 'comprehension' questions that required endorsement of a response in keeping with the interpretation provided in the scenario (i.e., a positive interpretation in positive trials and negative in negative trials). They received feedback on the accuracy of these answers, except on 'test' trials in which ambiguity had not been resolved[4].

Worry scenarios were adapted from Mathews and Mackintosh (2000), Hirsch *et al.* (2009), Hayes *et al.* (2010), and Grol, *et al.* (2018), while rumination scenarios were adapted from Holmes, Mathews, Dalgleish, and Mackintosh (2006), Hertel *et al.* (2014), and Blackwell *et al.* (2015). Further scenarios were created by the authors, resulting in 500 unique worry-related and 500 unique rumination-related scenarios (see Krahé *et al*., 2016 for further details of materials).

Participants selected one of the offered RNT main themes (see above) at the start of each session; this determined the type of scenarios used and also the theme participants worried/ruminated about if they were in the CBM_RNT condition.

Control. In each session, control participants heard 50 ambiguous scenarios that remained unresolved. Half of the scenarios were followed by a Yes/No 'comprehension' question (as above) relating to the ambiguity of the scenario, but which was never followed by feedback, so allowing either interpretation without correction. The remainder was related to a factual element of the scenario, and these were followed by accuracy feedback. The 'test' trials from the CBM-I conditions were also included.

Interpretation Bias Measures

Scrambled sentences test (Wenzlaff & Bates, 1998; Wenzlaff & Bates, 2000). The scrambled sentences test involved participants using five of six presented words to form

grammatically correct sentences, which could either be of negative or positive valence. For example, "looks the future bright very dismal" could be unscrambled to form the sentence "the future looks very bright" (positive) or "the future looks very dismal" (negative). Participants were presented with 20 sentences to unscramble in five minutes, whilst holding in mind a string of six digits. Half the sentences related to worry themes and were generated by the authors, while half related to depressive rumination selected from Wenzlaff and Bates (1998) and Wenzlaff and Bates (2000). The number of positive sentences divided by the total number of grammatically correct sentences generated serves as an index of interpretation bias, with a higher index (scores range from 0 to 1) denoting a more positive interpretation bias. Two separate lists of 20 items were counterbalanced across participants over the two visits.

Recognition test (based on Mathews & Mackintosh, 2000). In the first part of this task, participants read 20 ambiguous scenarios, and completed word fragments of the final word (which did not resolve the ambiguity – see Krahe *et al*., 2016 for details) and answered comprehension questions. After all scenarios had been completed, participants were presented with the title of each scenario, followed by four statements. Two statements were consistent with resolution of ambiguity in the scenario in either a positive or negative way (targets), while the other two statements were again positive or negative but were not legitimate interpretations (foils). Participants rated how similar each statement was to the meaning of the original scenario, with greater similarity ratings for positive targets indicating a more positive interpretation of that scenario (and similarly, greater similarity rating for negative targets indicating a more negative interpretation). Of the 20 scenarios, half related to worry (adapted from Mathews & Mackintosh, 2000, and Holmes *et al.*, 2006) and half related to depressive rumination (created by the authors – see Krahe *et al*., 2016 for examples). A recognition test index was computed for each participant by subtracting mean ratings for negative targets from mean ratings for positive targets. Thus, higher scores denoted greater similarity ratings to

positive vs. negative targets i.e., a more positive interpretation bias. Participants completed the recognition test before and after the first online session at the first visit, and again at the second (post-training) visit; hence, three separate sets of 20 items were generated, with set order counterbalanced across participants.

Worry and Rumination Measures

Breathing focus task (Hirsch et al., 2009; Hayes et al., 2010; Hirsch, Mathews, Lequertier, Perman, & Hayes, 2013). Participants focused on their breathing for five minutes and indicated at random cued intervals whether they were focusing on their breathing or experiencing a thought intrusion. They categorized thought intrusions as negative, positive, or neutral, and provided brief summaries of content. They then engaged in worry (GAD group) or rumination (MDD group) about a current worry/rumination topic for five minutes, followed by another five-minute breathing focus period, with sampling as before. After each breathing focus period, participants were reminded of their summaries of thought intrusion in turn, and gave expanded descriptions of the thoughts experienced at the time of sampling, which were audio-recorded for later categorization as negative, positive, or neutral by an assessor who was blind to diagnostic group, condition, and breathing phase (pre- vs. post-period of worry/rumination). Another rater categorized intrusions from 25% of participants and assessors had excellent agreement (ICC = .96, 95% CIs =.95, .97).

Standardized self-report questionnaires. See Krahe *et al*. (2016) for full details[4]. Trait worry levels were assessed using the Penn State Worry Questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990; Cronbach's α = .79 at baseline in the present sample). Trait rumination was measured using the Ruminative Response Scale (RRS; Nolen-Hoeksema & Morrow, 1991; Cronbach's α = .89).

[4]Participants also completed a novel 15-item 'RNT questionnaire' (RNTQ) that was designed to assess a range of different potential aspects of RNT and its consequences. This measure is under development and has not yet been validated and results are thus reported in Additional Supplementary Materials.

Anxiety and depression symptoms. Depressive symptoms were assessed with Patient Health Questionnaire 9 (PHQ-9; Kroenke & Spitzer, 2002; Cronbach's α = .73) and anxiety symptoms using the Generalized Anxiety Disorder 7-item scale (GAD-7; Spitzer *et al.*, 2006; Cronbach's α = .71).

Procedure

Participants completed questionnaires (PSWQ, RRS, PHQ-9, GAD-7, RNTQ) online within 24 hours prior to the first study visit. At the first visit, participants provided informed consent and were randomized by diagnostic group to one of the three conditions: 1) CBM_RNT, 2) CBM_STAND or 3) CONTROL (see *Experimental conditions* section above for details), and completed the scrambled sentences test, recognition test, and breathing focus task[5]. Then, participants were given a brief study rationale for the online sessions and completed expectancy ratings[6] before completing the first online session on the study website[7]. Each online session included the RNT induction (CBM_RNT) or neutral task (CBM_STAND and CONTROL conditions), followed by 50 scenarios (see above), with a short break after 25 scenarios. Following the first online session, participants completed the recognition test again.[8] Lastly, participants were presented with instructions for completing the nine online sessions at home over the course of the next month.

Participants then completed nine online sessions. We encouraged them to complete three sessions per week, with the expectation that there would be some slippage and thus we allowed up to one month in total. They were required to complete at least 8 online sessions

[5]Participants also completed the classic and emotional Stroop task during both experimental sessions. These were included for comparison with future experiments and did not relate to the current study aims. Results are presented in Additional Supplementary Materials.

[6]Results for expectancy and acceptability ratings (acceptability ratings were obtained at visit 2) are reported in Additional Supplementary Materials; participants expected the program to be moderately logical and useful, and indicated that the conditions (including the active control condition) were similar in this respect prior to training. After completing the program, participants in the CBM-I conditions reported the program to have been more useful than did those in the control condition.

[7] Participants were blind to their condition; experimenters were not blind to participants' condition, since they guided participants though the first online session on the website, which differed by condition.

[8]The findings for within-visit change on the RT are reported in Additional Supplementary Materials.

within one month with the final session no later than the day before their second (post-training) visit[9]. Researchers monitored adherence to the online sessions using the online platform. They kept in touch with participants by using participants' preferred method of contact (email, phone, SMS) to facilitate engagement and trouble shoot issues, and encourage participants to catch up with missed sessions.

Up to 24 hours before returning for their second (post-training) visit, participants completed PSWQ, RRS, PHQ-9, GAD-7, RNTQ questionnaires, and an "adverse events form" for the period since the first session[11]. Participants also completed acceptability, assimilation and imagination ratings (in reference to the home-based sessions[10]), and then completed the scrambled sentences test, recognition test, and breathing focus task, before being debriefed in general about the study. One month after their second study visit, participants completed the PSWQ, RRS, PHQ-9, GAD-7, RNTQ questionnaires, and the "adverse events form"[11]. Participants received £130 ($170) for their participation in the study. The study was approved by the ethics committee of the authors' university. Recruitment and testing commenced in January 2016 and final follow-up data was collected in January 2017.

Plan of Statistical Analyses

Statistical analyses were carried out in Stata 14 (StataCorp, 2015). Only participants who had completed at least 8 of the 10 online sessions were viewed as 'completers' and included in analyses.

Assessing the impact of multi-session CBM-I (Hypothesis 1) and modulatory role of RNT prior to CBM-I (Hypothesis 2).

[9]The average time between visits 1 and 2 was M = 22.76 days (SD = 2.82), indicating that participants completed the online sessions in just over three weeks; 74% of participants completed all 10 online sessions, 20% completed 9 sessions, and 6% completed 8 sessions.
[10]Findings are presented in Additional Supplementary Materials.
[11] The average time between visit 2 and completing the follow-up questionnaire was M = 31.93 days (SD = 5.49), indicating that the follow-up period was one month, as intended.

To assess the impact of multi-session CBM-I on our outcome measures (Hypothesis 1), we first compared both CBM-I conditions (combined) to the active control condition. To then investigate the potential modulatory role of prior RNT (Hypothesis 2), we contrasted only the two CBM-I conditions (CBM_RNT vs. CBM_STAND) and did not include the control condition. Our RNT manipulation check confirmed that the two CBM-I conditions differed as intended, i.e., that we successfully induced RNT prior to CBM-I in the CBM_RNT condition, by comparing self-report ratings of worry, rumination, anxiety, and depression (averaged across online sessions) immediately after RNT induction (CBM_RNT condition) vs. the neutral filler task (CBM_STAND) using a MANOVA[12]. Below, we outline the analytic strategy for testing Hypotheses 1 and 2 for each of our outcome measures.

Interpretation bias. For measures of interpretation bias, we collapsed analyses across diagnostic groups and conducted regression analyses (with bootstrapped standard errors in the case of non-normally distributed data) with mean score at the (post-training) second visit as the outcome variable and condition (combined CBM-I vs. control to address Hypothesis 1; CBM_RNT vs. CBM_STAND to address Hypothesis 2) as the predictor variable, and controlled for scores at the first visit (i.e., baseline scores). Establishing that our CBM-I training was successful in promoting a more positive interpretation bias was an important pre-requisite for assessing its consequent impact in terms of RNT and mood.

Repetitive negative thinking. For the self-report measures of worry and rumination, we again collapsed across diagnostic groups to see the overall impact of training. For worry and rumination separately, we specified multi-level regression models with mean questionnaire score as the outcome variables and condition (see above), post-training time

[12]The two CBM-I conditions (with and without prior RNT) differed on worry, rumination, anxiety and depression ratings immediately post RNT induction vs. neutral filler task (averaged across all online sessions), $F(4, 82) = 23.95$, $p < .001$, Wilk's lamda = .461. All ratings were higher in the CBM_RNT than CBM_STAND condition, as expected (see Integral Supplementary Table 1 for follow-up regression analyses and marginal means, also separately by group).

point (visit 2 and follow up) as predictor variables, and controlled for mean score at visit 1. A random effect was included to account for the repeated assessment of the outcome variable within individuals. Subsequently, we conducted planned subgroup analyses stratifying by diagnostic groups by specifying multi-group models (see Krahe *et al*., 2016) to examine whether there were any diagnosis-specific effects.

The breathing focus task included a further between-subjects factor of assessor type (participant self-report vs. independent assessor) and a repeated-measures factor of breathing focus period (pre-/post- induction of RNT). We specified a multi-level regression model with mean number of negative thought intrusions as the outcome variable and condition (as above), visit (visit 1, visit 2), assessor type (participant self-report vs. independent assessor), and breathing focus period (pre-/post- induction of RNT) as predictor variables. A random effect accounted for the repeated assessment of the outcome variable within individuals.

Anxiety and depression symptoms. For measures of anxiety and depression symptoms, we followed the same analytic strategy as that described for measures of worry and rumination above and report reliable change scores in Integral Supplementary Materials.

Assessing whether effects of CBM-I on worry, rumination, anxiety and depression were partially mediated by interpretation bias (Hypothesis 3).

To test Hypothesis 3, we examined whether change in interpretation bias partially mediated the effects of training on outcome measures at follow up using structural equation modelling following the product of coefficients approach (seen to be superior to the approach advocated by Baron & Kenny, 1986; Iacobucci, Saldanha, & Deng, 2007). In particular, we specified models to test whether interpretation bias (analyses run separately for SST and RT) mediated the effects of condition (combined CBM-I vs. control) on worry, rumination, anxiety, and depression scores at 1-month follow up (controlling for bias scores and symptom

scores at visit 1). Bootstrapping with 1000 replications estimated the standard errors. The proportion of the effect explained by the indirect path was calculated. The analyses were run across diagnostic groups.

Results

Descriptive Statistics

Descriptive statistics for outcome measures by condition and time point across diagnostic groups are presented in Table 2.

[INSERT TABLE 2 HERE]

Effects of Multi-Session CBM-I Training Compared to an Active Control Condition

Interpretation bias.

On the *Scrambled Sentences Test*, one participant was excluded from analyses for failing to complete any sentences in a grammatically correct fashion, leaving 130 participants for this analysis. Across diagnostic groups, the regression analysis showed that condition (combined CBM-I vs. control) was significantly associated with positivity index at visit 2 (Hedges' g = .33; see Table 2 for descriptive statistics and model results). As expected, participants in the combined CBM-I condition made more positive interpretations compared to the control condition following the 3-week online program. On the *Recognition Test*, one participant was excluded for not completing the task correctly, leaving 130 participants for this analysis. Across diagnostic groups, condition (combined CBM-I vs. control) was significantly associated with recognition test index at visit 2 (Hedges' g = .37; see Table 2). Participants made more positive (vs. negative) interpretations in the combined CBM-I compared to the control condition, as expected. Thus, findings on both measures of interpretation bias supported Hypothesis 1: interpretation bias was more positive following

CBM-I training compared to the active control condition, though the size of the effect was small to moderate. Having established change in interpretation bias, its impact on RNT and symptoms of anxiety and depression could now be assessed.

Levels of worry and rumination.

Self-report measures of worry (PSWQ) and rumination (RRS).

Seven participants were excluded from analyses of all the self-report questionnaires because they reported taking up psychological treatment between visit 2 and follow up (n = 4) or because they failed to complete the follow-up questionnaire (n = 3), leaving 124 participants for these multi-level regression analyses.

PSWQ. Across diagnostic groups, CBM-I was associated with significantly lower worry scores than the control condition at follow up but not visit 2 (see Table 2; Hedges' g_{visit2} = .15, Hedges' $g_{followup}$ = .29). A subgroup analysis examining whether the effect of condition was diagnosis specific revealed that the effect of CBM-I (vs. control) on worry was significant for the GAD group at follow up (Hedges' g = .30; see Table 3 for model results for subgroup analyses and descriptive statistics by group); the effect size for the MDD group was smaller and non-significant (Hedges' g = .12). Thus, at follow up, worry scores were lower following CBM-I than the control condition across diagnostic groups, supporting Hypothesis 1. This effect appeared to be most evident in the GAD group.

[INSERT TABLE 3 HERE]

RRS. Across diagnostic groups, CBM-I was significantly associated with lower rumination scores than the control condition at both post-training time points (see Table 2; Hedges' g_{visit2} = .31, Hedges' $g_{followup}$ = .51), supporting Hypothesis 1. Additionally, a subgroup analysis revealed that the effect of CBM-I on rumination was significant for the

MDD group at visit 2 (Hedges' g = .39) and the GAD group at follow up (Hedges' g = .39), although the effect size for the MDD group at follow up was of similar magnitude (Hedges' g = .36; see Table 3). Thus, post-training rumination scores were lower in the CBM-I than the control condition in both diagnostic groups, supporting Hypothesis 1.

Breathing focus task. For both self-reported and assessor-rated negative intrusions, condition was not significantly associated with number of negative thought intrusions post training, either before (Hedges' $g_{selfreport}$ = .24, Hedges' $g_{assessor}$ = .24) or after (Hedges' $g_{selfreport}$ = .29, Hedges' $g_{assessor}$ = .21) the RNT induction period (see Table 2). Thus, CBM-I training did not differ from the control condition, contrary to Hypothesis 1.

Taken together, there is sufficient evidence to conclude that effects of CBM-I training (vs. control) transferred to self-reported levels of worry and rumination, but not to a more behavioral measure of negative thought intrusions, providing partial support for Hypothesis 1.

Self-reported levels of anxiety (GAD-7) and depression (PHQ-9).

GAD-7. Across diagnostic groups, CBM-I was not associated with lower anxiety scores than the control condition at either time point, although the estimated effects at follow up were in the expected direction (see Table 2; Hedges' g_{visit2} = .04, Hedges' $g_{followup}$ = .27). A subgroup analysis examining whether the effect of condition was diagnosis specific revealed that the effect of CBM-I (vs. control) on anxiety was significant for the GAD group at follow up (Hedges' g = .56, see Table 3); the effect size for the MDD group at follow up was in the opposite direction (Hedges' g = -.16). At follow up, anxiety scores were lower in the CBM-I than the control condition in the GAD group only, partially supporting Hypothesis 1.

PHQ-9. Across diagnostic groups, CBM-I was significantly associated with lower depression scores than the control condition at both post-training time points (see Table 2; Hedges' g_{visit2} = .39, Hedges' $g_{followup}$ = .35), supporting Hypothesis 1. Additionally, a

subgroup analysis revealed that the effect of CBM-I on depression was significant for the MDD group at visit 2 (Hedges' $g = .48$), and, though not significant, effect sizes for the GAD group at visit 2 and follow up were of similar magnitude (Hedges' $g_{visit2} = .31$, Hedges' $g_{followup} = .30$; see Table 3). Thus, post-training depression scores were lower in the CBM-I than the control condition across both groups, supporting Hypothesis 1.

Overall, the findings provide support for Hypothesis 1. Across groups, CBM-I – with and without prior RNT activation – promoted a more positive interpretation bias and reduced levels of worry and rumination (though on self-report measures only), and levels of depression relative to the active control condition. Beneficial effects of CBM-I on levels of anxiety appeared specific to the GAD group.

Potential Modulatory Role of Prior RNT on the Effects of CBM-I Training

Results for the comparison of the two CBM-I conditions on outcome measures are presented in Integral Supplementary Tables 2 and 3. Contrary to Hypothesis 2, we found no significant differences between CBM_RNT and CBM_STAND conditions on any of the outcome measures. Thus, there was insufficient evidence to conclude that engaging in RNT prior to completing CBM-I training modulated the effects of CBM-I on outcome measures.

Assessing Whether Interpretation Bias Post Training Partially Mediated Effects of CBM-I on Outcomes at Follow Up

To test Hypothesis 3, we examined whether change in interpretation bias at visit 2 mediated the effect of condition (combined CBM-I vs. control) on outcome measures at 1-month follow up, using structural equation modelling. Interpretation bias as measured by SST (but not RT) partially mediated the effect of CBM-I on worry, rumination, anxiety, and depression (see Table 4 for all results). The indirect path from condition to PSWQ, RRS, and GAD-7 scores via SST visit 2 score were significant: the proportion of the effect mediated

was .36 for PSWQ, .24 for RRS, and .31 for GAD-7. The indirect path from condition to PHQ-9 score via SST visit 2 score approached significance (p = .054), and the proportion of the effect mediated was .23. Thus, a quarter to a third of the effect of condition on follow-up questionnaire scores was mediated by level of interpretation bias (as measured by the SST) after completing the online program, providing partial support for Hypothesis 3[13].

[INSERT TABLE 4 HERE]

Discussion

Overall, consistent with our main hypothesis, online training designed to reduce negative interpretation bias was effective in promoting a more positive interpretation bias and in reducing the targeted form of RNT and depressed mood across both diagnostic groups (GAD and depression). There was no evidence for CBM-I effects on a more behavioral measure of RNT, i.e., the number of negative thought intrusions reported immediately after training, and effects varied according to diagnosis for anxiety symptoms, as will be discussed later. However, the overall pattern of results supported the effectiveness of the current training in reducing worry and rumination, with related improvements in depression and anxiety.

It is not clear why this finding differs from some other reports of failure to obtain persisting change in clinical symptoms following internet delivered CBM (e.g., Blackwell *et al.*, 2015; Carlbring *et al.*, 2012). Possible reasons include differences in the type of training used, the bias that was targeted, the clinical status of participants, and the control condition used, perhaps in different combinations across studies. For example, Carlbring *et al.* (2012)

[13] In addition, to demonstrate that change in cognitive bias was related to change in symptoms, we computed bivariate correlations for change scores (visit 2 minus visit 1, as we did not have follow-up data for bias measures) for the combined CBM-I condition. Bias change on the SST (an increase in positivity) was significantly correlated with all measures of symptom change (a reduction in scores; Pearson's r's ranged from -.26 to -.38, $p < .05$), while – mirroring the mediation results – bias change on the RT was not correlated with symptom change, apart from one significant correlation between RT change and change on the GAD-7 (r = -.24, $p < .05$).

studied the effect of attention training on social anxiety, and suggested that their null results may have been due to the lack of emotional activation during home practice. The improved outcome in a follow-up study (Kuckertz *et al.*, 2014) was thus attributed to the requirement of engaging in social activity prior to training sessions at home. However, this form of activation necessarily involves exposure to feared situations, which may itself be sufficient to account for the greater improvement seen in their second (follow-up) study.

Some other studies (e.g., Blackwell *et al.*, 2015) used similar interpretation training methods, but used a control condition in which ambiguous events were resolved in a negative manner half the time and positively in the other half. As noted by these authors, relative to their own short-term experimental studies, this control condition resulted in greater reduction in depressed mood over longer training, perhaps related to participants learning that ambiguity can be resolved in different (and sometimes positive) ways. Whatever the explanation, this unexpected improvement in the 50:50 control condition in Blackwell *et al.* (2015) may have partly accounted for the failure to show superiority of interpretation training. As noted in the earlier discussion of Blackwell *et al.* (2017), to address our research question, we elected to use a control condition that left ambiguous training items unresolved, and this did not result in marked changes in bias over time, thus allowing a more sensitive assessment of the causal role of negative interpretations in RNT.

An additional hypothesis in the present study, namely that priming negative concerns via RNT might modulate training effects, was not supported. This latter hypothesis was based (in part) on prior animal and human research on memory reconsolidation, showing that activation of old fear memories shortly prior to incompatible re-learning renders the original memories more malleable (Jaffe, 2013). Our extrapolation assumed that RNT is at least partly based on prior memories. However, the type of memory involved in cognitive bias may be more procedural in nature (e.g., in the form of a habitual production rule), rather than episodic

as in memory for unambiguous fear-inducing events. Furthermore, some evidence suggests that emotional memories are more resistant to change in highly anxious individuals (Kindt & Soeter, 2013). For this or other procedural reasons (e.g., activation was immediately followed by CBM-I rather than after a delay, that may be critical for reconsolidation), our assumption that RNT activation could enhance CBM-I was not supported. Furthermore, given the lack of a modulatory role of prior RNT, the suggestion that activation via RNT may have the counter-productive effect of interfering with the cognitive resources needed for relearning was also not supported. It is conceivable that the two (beneficial vs. detrimental) effects cancelled each other out, but we see no evidence to support this somewhat implausible idea. Future research could investigate whether more prolonged RNT activation and a rest period between RNT activation and CBM (more closely mirroring the conditions thought critical for the intended reconsolidation effect) are necessary for beneficial effects of activation in CBM-I to manifest.

As noted earlier, although the overall pattern of results held over both diagnostic groups, there were some minor differences. It is perhaps not surprising that bias modification would have greater effects on emotional responses characteristic of the given clinical condition. Specifically, a training-related reduction in anxiety was found at follow up only in participants diagnosed with GAD. It may be that promoting more positive interpretations related to rumination did not transfer to reductions in anxiety, perhaps because rumination is less anxiogenic than worry. Nevertheless, by follow-up, reductions in worry were found across diagnostic groups (although most evident in GAD), and changes in rumination were of similar magnitude in both GAD and depression at follow up. These findings suggest that training interpretations related to one form of RNT can transfer to reductions in the other form of RNT, consistent with the possibility that interpretation bias is a common cognitive mechanism underlying both worry and rumination.

However, while reductions in rumination and depression were evident across diagnostic groups at both post-training time points, reductions in worry and anxiety (the latter confined to the GAD group) were significant at 1-month follow up only. It is possible that rumination and corresponding low mood may be more changeable than worry and anxiety; indeed, a diagnosis of depression is currently based on the last month, while a diagnosis of GAD is based on the last six months (American Psychiatric Association, 2013). Furthermore, while natural remission is common in depression, GAD tends to follow a chronic time course (Kessler, Keller, & Wittchen, 2001). Thus, although interpretation bias seems to play a role in both worry and rumination, the time course and route by which it influences anxiety and depression may differ between the two clinical disorders.

A related limitation is that the greater changes in bias following training were not reflected in similar differential reductions of negative intrusive thoughts during the breathing focus task. For participants with GAD, training effects on worry and anxiety were not apparent until follow-up, rather than earlier, at the time at which intrusive thoughts were reassessed. Future research could therefore involve later assessment of negative intrusions to evaluate whether these change at follow up operate in concert with reported reductions in worry. In depression, it may be that interpretation bias is less involved in promoting intrusions, as opposed to persistence of RNT, which was not assessed in this task but may be particularly relevant to rumination. Testing these ideas would require a different method, assessing the duration and negativity of RNT episodes, instead of only intrusion frequency, and at follow up as well as earlier. Mediation analyses suggested that interpretation bias as measured by the Scrambled Sentence Task (SST) partially mediated changes in both worry and rumination, as well as anxiety (with a similar but non-significant trend in depression symptoms). In contrast, bias as measured in the recognition test was not found to mediate outcome, nor to correlate with changes in RNT. There are a number of differences between

these two measures of interpretation that might contribute to this differential mediation. The recognition test is much closer in format to training items than is the SST, while the latter involves more active selection and rejection of simultaneously presented alternatives. Conceivably, these differences result in the SST more closely representing the processes required to select positive and reject competing negative interpretations in real life situations, thus accounting for more evidence of mediation. This idea could be investigated in future studies by building in systematic variations to the type of interpretation tests used, with the hypothesis that versions having more novel format, and demanding more active selection, are more likely to mediate generalization to RNT and related symptoms.

A final limitation was that we included participants with either GAD or depression, but not with a dual diagnosis of both disorders. This decision was in part related to difficulties in assigning participants to training materials (whether they should receive worry or rumination materials or a mix of both). However, given the high comorbidity of GAD and depression and our finding that training related to one form of RNT also transferred to the other form, future research could include participants with both GAD and depression, which would increase both the diversity of the sample and the generalizability of the findings. In a related vein, future research could seek to extend the 1-month follow up period to examine whether changes in symptoms are maintained over longer periods.

Despite the above limitations, we believe the present study also has several strengths. Including participants with diagnoses of GAD or Depression meant that participants had pathological levels of worry and rumination at study baseline. Interestingly, the drop out from the study was low, despite it involving many home-based online sessions, indicating that interpretation bias training is an acceptable form of intervention, even for those with depression, where engagement can sometimes be particularly challenging. Furthermore, the present study is the first to examine the effects of multi-session CBM-I on worry and

rumination within the same experimental study. The inclusion of two separate measures of interpretation bias, with effects of training evident for both, is a further strength. CBM was originally developed as an experimental tool to assess the causal role of cognitive biases in maintaining psychopathology. While other research in the field sets out to examine CBM's potential therapeutic application, a strength of our experimental research is that is focuses on understanding underlying mechanisms of RNT and as such was not intended to be a clinical trial. We believe that this experimental approach to enhancing understanding of mechanisms and moderators will facilitate the development of improved interventions in the future.

Given our main aim of understanding the causal role of interpretation bias in RNT, we conclude by briefly considering how negative interpretations may operate to maintain RNT. When individuals encounter ambiguous information, a tendency to generate negative interpretations will enhance the perceived threat and may thus serve to initiate an episode of RNT. Once begun, interpretation bias operating on the content of RNT itself is likely to direct thoughts to the worst possibilities, enhancing the negativity of RNT content (see Hirsch & Mathews, 2012). Furthermore, other cognitive biases may then also influence these streams of negative thinking, for example, a tendency to think in more verbal abstract ways (Hirsch, Hayes, Mathews, Perman, & Borkovec, 2012), as well as a depletion of the cognitive resources required to shift away from negative content (Stefanopoulou *et al.*, 2014), both of which impede stopping RNT and shifting to alternative content.

Clearly, future research needs to address the aim of enhancing training methods and unequivocally establishing the causal role of emotional processing biases. For example, rather than activation via RNT, another route to enhancing training effects may be to identify and remediate individual obstacles to bias malleability, such as the failure to engage with or identify with the material used in training (see Standage, Harris, & Fox, 2014). Nevertheless, in contrast to recent failures to demonstrate improvements following internet-delivered CBM,

we have shown that self-reported RNT and negative mood can be reduced by home-based CBM-I. This is consistent with a causal role for negative interpretation bias in both worry and rumination, suggesting a useful role for CBM-I within treatments for anxiety and depression.

Acknowledgments

We are grateful to our grant steering committee members with lived experience of generalized anxiety disorder and/or depression and Emese Nagy for helping with the development of materials for the study, to Emma Drinkwater-James for audio recording the study scenarios, to Amy Steele and Josh Graham for their help with data collection and data entry, to Alice Kerr for independently rating a subset of the SCIDs, to Jasmine Vickers for providing the assessor ratings of thought intrusions in the breathing focus task, and Sarah Illsley for independently providing assessor ratings for a subset of participants' breathing focus thought intrusions. We would also like to thank Prof Lusia Stopa, Dr Nick Grey and steering committee members with lived experience of GAD and/or Depression for their helpful contributions during grant steering committee meetings.

This work was supported by an MQ: Transforming Mental Health PsyIMPACT grant (number MQ14PP_84) to CH. CH receives salary support from the National Institute for Health Research (NIHR), Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of MQ: Transforming Mental Health, NHS, the National Institute for Health Research (NIHR), or the Department of Health.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. (5th ed.). Arlington, VA: American Psychiatric Publishing.

Amir, N., & Taylor, C. T. (2012). Interpretation training in individuals with generalized social anxiety disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 80*, 497-511.

Anderson, K. G., Dugas, M. J., Koerner, N., Radomsky, A. S., Savard, P., & Turcotte, J. (2012). Interpretive style and intolerance of uncertainty in individuals with anxiety disorders: A focus on generalized anxiety disorder. *Journal of anxiety disorders, 26*, 823-832.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Berna, C., Lang, T. J., Goodwin, G. M., & Holmes, E. A. (2011). Developing a measure of interpretation bias for depressed mood: An ambiguous scenarios test. *Personality and Individual Differences, 51*, 349-354.

Blackwell, S. E., Browning, M., Mathews, A., Pictet, A., Welch, J., Davies, J., et al. (2015). Positive imagery- based cognitive bias modification as a web-based treatment tool for depressed adults: A randomized controlled trial. *Clinical Psychological Science, 3*, 91-111.

Blackwell, S. E., Woud, M. L., & MacLeod, C. (2017). A question of control? Examining the role of control conditions in experimental psychopathology using the example of cognitive bias modification research. *The Spanish Journal of Psychology, 20*.

Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy, 5*, 51-62.

Carlbring, P., Apelstrand, M., Sehlin, H., Amir, N., Rousseau, A., Hofmann, S. G., et al. (2012). Internet-delivered attention bias modification training in individuals with social anxiety disorder-a double blind randomized controlled trial. *BMC Psychiatry, 12*, 66.

Cohen, N., Mor, N., & Henik, A. (2015). Linking executive control and emotional response: A training procedure to reduce rumination. *Clinical Psychological Science, 3*, 15-25.

Drost, J., van der Does, W., van Hemert, A. M., Penninx, B. W. J. H., & Spinhoven, P. (2014). Repetitive negative thinking as a transdiagnostic factor in depression and anxiety: A conceptual replication. *Behaviour Research and Therapy, 63*, 177-183.

Everaert, J., Podina, I. R., & Koster, E. H. W. (2017). A comprehensive meta-analysis of interpretation biases in depression. *Clinical Psychology Review, 58*, 33-48.

Eysenck, M. W., Mogg, K., May, J., Richards, A., & Mathews, A. (1991). Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology, 100*, 144-150.

First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. (2015). *Structured clinical interview for dsm-5—research version (scid-5 for dsm-5, research version; scid-5-rv)*. Arlington, VA: American Psychiatric Association.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 245-262.

Grey, S., & Mathews, A. (2000). Effects of training on interpretation of emotional ambiguity. *The Quarterly Journal of Experimental Psychology: Section A, 53*, 1143-1162.

Grisham, J. R., Flower, K. N., Williams, A. D., & Moulds, M. L. (2011). Reappraisal and rumination during recall of a sad memory. [journal article]. *Cognitive Therapy and Research, 35*, 276-283.

Grol, M., Schwenzfeier, A., Stricker, J., Booth, C., Temple-McCune, A., Derakshan, N., et al. (2018). The worrying mind in control: An investigation of adaptive working memory training and cognitive bias modification in worry-prone individuals. *Behaviour Research and Therapy, 103, 1-11*.

Gyani, A., Shafran, R., Layard, R., & Clark, D. M. (2013). Enhancing recovery rates: Lessons from year one of iapt. *Behaviour Research and Therapy, 51*, 597-606.

Hayes, S., Hirsch, C., & Mathews, A. (2008). Restriction of working memory capacity during worry. *Journal of Abnormal Psychology, 117*, 712-717.

Hayes, S., Hirsch, C. R., Krebs, G., & Mathews, A. (2010). The effects of modifying interpretation bias on worry in generalized anxiety disorder. *Behaviour Research and Therapy, 48*, 171-178.

Hertel, P. T., & El-Messidi, L. (2006). Am I blue? Depressed mood and the consequences of self-focus for the interpretation and recall of ambiguous words. *Behavior Therapy, 37*, 259-268.

Hertel, P. T., & Mathews, A. (2011). Cognitive bias modification past perspectives, current findings, and future applications. *Perspectives on Psychological Science, 6*, 521-536.

Hertel, P. T., Mor, N., Ferrari, C., Hunt, O., & Agrawal, N. (2014). Looking on the dark side: Rumination and cognitive-bias modification. *Clinical Psychological Science, 2*, 714-726.

Hindash, A. H. C., & Rottenberg, J. A. (2017). Moving towards the benign: Automatic interpretation bias modification in dysphoria. *Behaviour Research and Therapy, 99*, 98-107.

Hirsch, C. R., Hayes, S., & Mathews, A. (2009). Looking on the bright side: Accessing benign meanings reduces worry. *Journal of Abnormal Psychology, 118*, 44-54.

Hirsch, C. R., Hayes, S., Mathews, A., Perman, G., & Borkovec, T. (2012). The extent and nature of imagery during worry and positive thinking in generalized anxiety disorder. *Journal of Abnormal Psychology, 121*, 238-243.

Hirsch, C. R., & Mathews, A. (2012). A cognitive model of pathological worry. *Behaviour Research and Therapy, 50*, 636-646.

Hirsch, C. R., Mathews, A., Lequertier, B., Perman, G., & Hayes, S. (2013). Characteristics of worry in generalized anxiety disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 44*, 388-395.

Hirsch, C. R., Meeten, F., Krahé, C., & Reeder, C. (2016). Resolving ambiguity in emotional disorders: The nature and role of interpretation biases. *Annual Review of Clinical Psychology, 12*, 281-305.

Holmes, E. A., & Mathews, A. (2005). Mental imagery and emotion: A special relationship? *Emotion, 5*, 489.

Holmes, E. A., Mathews, A., Dalgleish, T., & Mackintosh, B. (2006). Positive interpretation training: Effects of mental imagery versus verbal training on positive mood. *Behavior Therapy, 37*, 237-247.

Hoppitt, L., Mathews, A., Yiend, J., & Mackintosh, B. (2010). Cognitive bias modification: The critical role of active training in modifying emotional responses. *Behavior Therapy, 41*, 73-81.

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology, 17*, 139-153.

Jaffe, E. (2013). Building a fearless mind. *APS Observer*.

Kessler, R. C., Keller, M. B., & Wittchen, H.-U. (2001). The epidemiology of generalized anxiety disorder. *Psychiatric Clinics of North America, 24*, 19-39.

Kindt, M., & Soeter, M. (2013). Reconsolidation in a human fear conditioning study: A test of extinction as updating mechanism. *Biological Psychology, 92*, 43-50.

Krahé, C., Mathews, A., Whyte, J., & Hirsch, C. R. (2016). Cognitive bias modification for interpretation with and without prior repetitive negative thinking to reduce worry and rumination in generalised anxiety disorder and depression: Protocol for a multisession experimental study with an active control condition. *BMJ Open, 6*, e013404.

Krahé, C., Whyte, J., Bridge, L., Loizou, S., & Hirsch, C. R. (submitted). *Interpretation bias predicts worry and rumination in generalized anxiety disorder and depression.*

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals, 32*, 509-515.

Kuckertz, J. M., Gildebrant, E., Liliequist, B., Karlström, P., Väppling, C., Bodlund, O., et al. (2014). Moderation and mediation of the effect of attention training in social anxiety disorder. *Behaviour Research and Therapy, 53*, 30-40.

Lang, T. J., Blackwell, S. E., Harmer, C. J., Davison, P., & Holmes, E. A. (2012). Cognitive bias modification using mental imagery for depression: Developing a novel computerized intervention to change negative thinking styles. *European Journal of Personality, 26*, 145-157.

Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*, 602-615.

Mathews, A., Richards, A., & Eysenck, M. (1989). Interpretation of homophones related to threat in anxiety states. *Journal of Abnormal Psychology, 98*, 31-34.

McMillan, D., Gilbody, S., & Richards, D. (2010). Defining successful treatment outcome in depression using the PHQ-9: A comparison of methods. *Journal of Affective Disorders, 127*, 122-129.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the penn state worry questionnaire. *Behaviour Research and Therapy, 28*, 487-495.

Mogg, K., Baldwin, D. S., Brodrick, P., & Bradley, B. P. (2004). Effect of short-term ssri treatment on cognitive bias in generalised anxiety disorder. *Psychopharmacology, 176*, 466-470.

Mor, N., Hertel, P., Ngo, T. A., Shachar, T., & Redak, S. (2014). Interpretation bias characterizes trait rumination. *Journal of Behavior Therapy and Experimental Psychiatry, 45*, 67-73.

Murphy, R., Hirsch, C. R., Mathews, A., Smith, K., & Clark, D. M. (2007). Facilitating a benign interpretation bias in a high socially anxious population. *Behaviour Research and Therapy, 45*, 1517-1529.

Nader, K., & Hardt, O. (2009). A single standard for memory: The case for reconsolidation. [10.1038/nrn2590]. *Nature Reviews Neuroscience, 10*, 224-234.

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 loma prieta earthquake. *Journal of Personality and Social Psychology, 61*, 115-121.

Nunn, J. D., Mathews, A., & Trower, P. (1997). Selective processing of concern-related information in depression. *British Journal of Clinical Psychology, 36*, 489-503.

Pictet, A., Jermann, F., & Ceschi, G. (2016). When less could be more: Investigating the effects of a brief internet-based imagery cognitive bias modification intervention in depression. *Behaviour Research and Therapy, 84*, 45-51.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The gad-7. *Archives of Internal Medicine, 166*, 1092-1097.

Standage, H., Harris, J., & Fox, E. (2014). The influence of social comparison on cognitive bias modification and emotional vulnerability. *Emotion, 14*, 170-179.

StataCorp. (2015). *Stata statistical software: Release 14*. College Station, TX: StataCorp LP.

Stefanopoulou, E., Hirsch, C. R., Hayes, S., Adlam, A., & Coker, S. (2014). Are attentional control resources reduced by worry in generalized anxiety disorder? *Journal of Abnormal Psychology, 123*, 330-335.

Suarez, L., & Bell-Dolan, D. (2001). The relationship of child worry to cognitive biases: Threat interpretation and likelihood of event occurrence. *Behavior Therapy*, *32*, 425-442.

Torkan, H., Blackwell, S. E., Holmes, E. A., Kalantari, M., Neshat-Doost, H. T., Maroufi, M., et al. (2014). Positive imagery cognitive bias modification in treatment-seeking patients with major depression in iran: A pilot study. *Cognitive therapy and research, 38*, 132-145.

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin, 134*, 163–206.

Wenzlaff, R. M., & Bates, D. E. (1998). Unmasking a cognitive vulnerability to depression: How lapses in mental control reveal depressive thinking. *Journal of Personality and Social Psychology, 75*, 1559-1571.

Wenzlaff, R. M., & Bates, D. E. (2000). The relative efficacy of concentration and suppression strategies of mental control. *Personality and Social Psychology Bulletin, 26*, 1200-1212.

Williams, A. D., Blackwell, S. E., Holmes, E. A., & Andrews, G. (2013). Positive imagery cognitive bias modification (cbm) and internet-based cognitive behavioural therapy (icbt) versus control cbm and icbt for depression: Study protocol for a parallel-group randomised controlled trial. *BMJ Open, 3*.

Williams, M. O., Mathews, A., & Hirsch, C. R. (2014). Verbal worry facilitates attention to threat in high-worriers. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*, 8-14.

Wisco, B. E., & Nolen-Hoeksema, S. (2010). Interpretation bias and depressive symptoms: The role of self-relevance. *Behaviour Research and Therapy*, *48*, 1113-1122.

Table 1. *Demographic Characteristics By Condition (Across Both Diagnostic Groups).*

		CBM_RNT	CBM_STAND	Control
Age - mean (SD)		29.76 (11.49)	29.47 (10.88)	30.59 (11.17)
Gender (F/M)		32/11	38/6	38/6
Nationality - N (%)	British	31 (72.09)	34 (77.27)	34 (77.27)
	Other European	7 (16.28)	6 (13.64)	7 (15.91)
	World	5 (11.63)	4 (9.09)	3 (6.82)
Highest level of education - N (%)	Secondary	13 (30.23)	8 (18.18)	12 (27.27)
	Bachelor	17 (39.53)	21 (47.73)	20 (45.45)
	Master	9 (20.93)	8 (18.18)	10 (22.73)
	Doctoral	0 (0)	0 (0)	2 (4.55)
	Other	4 (9.3)	7 (15.91)	0 (0)
Marital status - N (%)	Single, never married	35 (81.4)	32 (72.73)	30 (68.18)
	Married /domestic partnership	6 (13.95)	10 (22.73)	10 (22.73)
	Separated, divorced, widowed	2 (4.65)	2 (4.55)	4 (9.09)

Note. CBM_RNT = Cognitive Bias Modification for Interpretation with prior repetitive negative thinking; CBM_STAND = Cognitive Bias Modification for Interpretation without prior repetitive negative thinking

Table 2. *Descriptive Statistics And Model Results For Combined CBM-I Vs. Control Condition For All Outcome Measures Across Groups.*

			CBM-I			Control			Adjusted mean difference				
	Time point	Pre-/post RNT induction	*N*	Mean	*SD*	*N*	Mean	*SD*	*b*	*SE*	*p* value	95% CI lower	95% CI higher	Hedges' *g*
SST	Visit 1		87	0.44	0.18	43	0.50	0.20						
	Visit 2		87	0.57	0.20	43	0.54	0.23	-0.07	0.03	**0.035**	-0.13	0.00	0.33
RT	Visit 1		87	-0.20	0.72	43	-0.11	0.70						
	Visit 2		87	0.52	0.67	43	0.28	0.86	-0.28	0.12	**0.026**	-0.52	-0.03	0.37
PSWQ	Visit 1		86	67.55	6.90	38	64.82	8.45						
	Visit 2		86	64.56	8.45	38	63.68	8.48	1.26	1.41	0.372	-1.51	4.03	0.15
	Follow up		86	60.27	10.09	38	61.13	9.84	3.00	1.41	**0.034**	0.23	5.77	0.29
RRS	Visit 1		86	60.50	10.50	38	59.24	12.03						
	Visit 2		86	54.09	11.26	38	56.92	12.88	3.68	1.86	**0.048**	0.03	7.33	0.31
	Follow up		86	47.40	11.86	38	53.11	13.89	6.56	1.86	**0.000**	2.91	10.21	0.51
BFT-self	Visit 1	Pre	87	3.36	2.28	44	3.11	2.51	-0.24	0.41	0.554	-1.05	0.56	-0.10

report		Post	87	3.72	2.37	44	3.45	2.77	-0.27	0.41	0.511	-1.07	0.53	-0.11
	Visit 2	Pre	87	1.74	1.74	44	2.23	2.37	0.49	0.41	0.230	-0.31	1.29	0.24
		Post	87	2.03	1.93	44	2.70	2.77	0.67	0.41	0.102	-0.13	1.47	0.29
BFT-	Visit 1	Pre	87	2.18	2.01	41	2.34	2.54	0.10	0.42	0.809	-0.71	0.92	0.04
assessor		Post	87	3.07	2.32	40	3.25	2.89	0.09	0.42	0.825	-0.73	0.91	0.04
	Visit 2	Pre	85	1.27	1.55	44	1.75	2.15	0.44	0.41	0.286	-0.37	1.24	0.24
		Post	87	1.79	2.00	43	2.30	2.56	0.47	0.41	0.257	-0.34	1.27	0.21
GAD-7	Visit 1		86	13.55	3.53	38	13.89	3.10						
	Visit 2		86	11.07	4.69	38	11.42	4.10	0.17	0.87	0.844	-1.53	1.87	0.04
	Follow up		86	8.57	5.04	38	10.18	5.29	1.43	0.87	0.099	-0.27	3.14	0.27
PHQ-9	Visit 1		86	14.36	4.37	38	14.92	4.56						
	Visit 2		86	10.63	4.93	38	12.97	5.55	2.02	0.93	**0.030**	0.20	3.85	0.39
	Follow up		86	8.97	5.40	38	11.37	6.49	2.08	0.93	**0.026**	0.25	3.90	0.35

Note. CBM-I = Cognitive Bias Modification for Interpretation; RNT = Repetitive Negative Thinking; SST = Scrambled Sentences Test; RT = Recognition Test; BFT = Breathing focus task (self-reported intrusions only); PSWQ = Penn State Worry Questionnaire; RRS = Ruminative Response Scale; PHQ-9 = Patient Health Questionnaire (measure of depression); GAD-7 = Generalized Anxiety Disorder scale (measure of anxiety)

Table 3. *Descriptive Statistics And Model Results For Subgroup Analyses For Questionnaire Measures In Combined CBM-I Vs. Control Condition.*

			CBM-I			Control			Adjusted mean difference				
	Time point	Group	*N*	Mean	*SD*	*N*	Mean	*SD*	*b*	*SE*	*p* value	95% CI lower	95% CI higher	Hedges' *g*
PSWQ	Visit 1	MDD	42	65.52	7.07	18	62.28	9.74						
		GAD	44	69.48	6.22	20	67.10	6.54						
	Visit 2	MDD	42	63.76	7.38	18	61.33	9.65	-0.21	1.70	0.901	-3.55	3.12	-0.02
		GAD	44	65.32	9.39	20	65.80	6.84	2.63	1.77	0.137	-0.84	6.09	0.29
	Follow up	MDD	42	58.90	9.78	18	58.06	10.24	1.22	2.31	0.598	-3.32	5.76	0.12
		GAD	44	61.57	10.33	20	63.90	8.82	3.07	1.49	**0.040**	0.14	6.00	0.30
RRS	Visit 1	MDD	42	61.81	10.07	18	63.67	9.54						
		GAD	44	59.25	10.87	20	55.25	12.85						
	Visit 2	MDD	42	55.43	10.81	18	61.33	11.19	4.45	2.13	**0.037**	0.27	8.63	0.39
		GAD	44	52.82	11.66	20	52.95	13.26	2.69	2.61	0.301	-2.42	7.80	0.21
	Follow up	MDD	42	48.19	11.66	18	55.11	14.51	4.66	2.99	0.119	-1.2	10.53	0.36
		GAD	44	46.64	12.14	20	51.30	13.41	5.14	2.10	**0.014**	1.02	9.25	0.39

GAD-7	Visit 1	MDD	42	13.07	3.85	18	13.17	2.55					
		GAD	44	14.00	3.18	20	14.55	3.46					
	Visit 2	MDD	42	10.52	4.44	18	11.72	3.94	1.14	1.03	0.267	-0.87	3.15	0.25
		GAD	44	11.59	4.91	20	11.15	4.33	-0.66	1.21	0.588	-3.04	1.72	-0.13
	Follow up	MDD	42	8.76	4.45	18	8.72	5.39	-0.79	1.07	0.461	-2.89	1.31	-0.16
		GAD	44	8.39	5.58	20	11.50	4.96	3.17	1.15	**0.006**	0.92	5.43	0.56
PHQ-9	Visit 1	MDD	42	15.6	4.70	18	16.94	3.57					
		GAD	44	13.18	3.71	20	13.10	4.66					
	Visit 2	MDD	42	11.17	4.61	18	14.33	5.65	2.45	1.22	**0.044**	0.07	4.83	0.48
		GAD	44	10.11	5.23	20	11.75	5.31	1.69	1.23	0.169	-0.72	4.09	0.31
	Follow up	MDD	42	9.17	5.06	18	11.22	7.22	-0.32	1.24	0.798	-2.74	2.11	-0.05
		GAD	44	8.77	5.75	20	11.50	5.95	1.83	1.20	0.128	-0.53	4.18	0.30

Note. CBM-I = Cognitive Bias Modification for Interpretation; PSWQ = Penn State Worry Questionnaire; RRS = Ruminative Response Scale; PHQ-9 = Patient Health Questionnaire (measure of depression); GAD-7 = Generalized Anxiety Disorder scale (measure of anxiety); GAD = Generalized Anxiety Disorder; MDD = Depression

Table 4. *Mediation Analyses Assessing Whether Interpretation Bias Post Training Mediates Effects Of CBM-I Vs. Control On Questionnaire Measures Of Repetitive Negative Thinking And Mood At Follow-up.*

		Effect	*b*	bootstrapped SE	*p* value
SST	PSWQ	Direct	2.16	1.43	0.131
		Indirect	1.23	0.60	0.041
		Total	3.39	1.46	0.020
		Proportion mediated	0.36		
	RRS	Direct	5.55	1.98	0.005
		Indirect	1.73	0.78	0.026
		Total	7.28	2.11	0.001
		Proportion mediated	0.24		
	GAD-7	Direct	1.38	0.77	0.075
		Indirect	0.62	0.31	0.044
		Total	2.00	0.79	0.012
		Proportion mediated	0.31		
	PHQ-9	Direct	2.07	0.89	0.020
		Indirect	0.63	0.33	0.054
		Total	2.71	0.94	0.004
		Proportion mediated	0.23		
RT	PSWQ	Direct	1.87	1.54	0.227
		Indirect	1.05	0.64	0.099
		Total	2.92	1.59	0.067
		Proportion mediated	0.36		

RRS	Direct	4.85	2.13	0.023
	Indirect	1.35	0.87	0.121
	Total	6.20	2.21	0.005
	Proportion mediated	0.22		
GAD-7	Direct	0.80	0.90	0.377
	Indirect	0.55	0.32	0.093
	Total	1.34	0.94	0.154
	Proportion mediated	0.41		
PHQ-9	Direct	1.53	1.03	0.135
	Indirect	0.44	0.31	0.157
	Total	1.98	1.07	0.065
	Proportion mediated	0.22		

Note. SST = Scrambled Sentences Test; RT = Recognition Test; PSWQ = Penn State Worry Questionnaire; RRS = Ruminative Response Scale; PHQ-9 = Patient Health Questionnaire (measure of depression); GAD-7 = Generalized Anxiety Disorder scale (measure of anxiety)

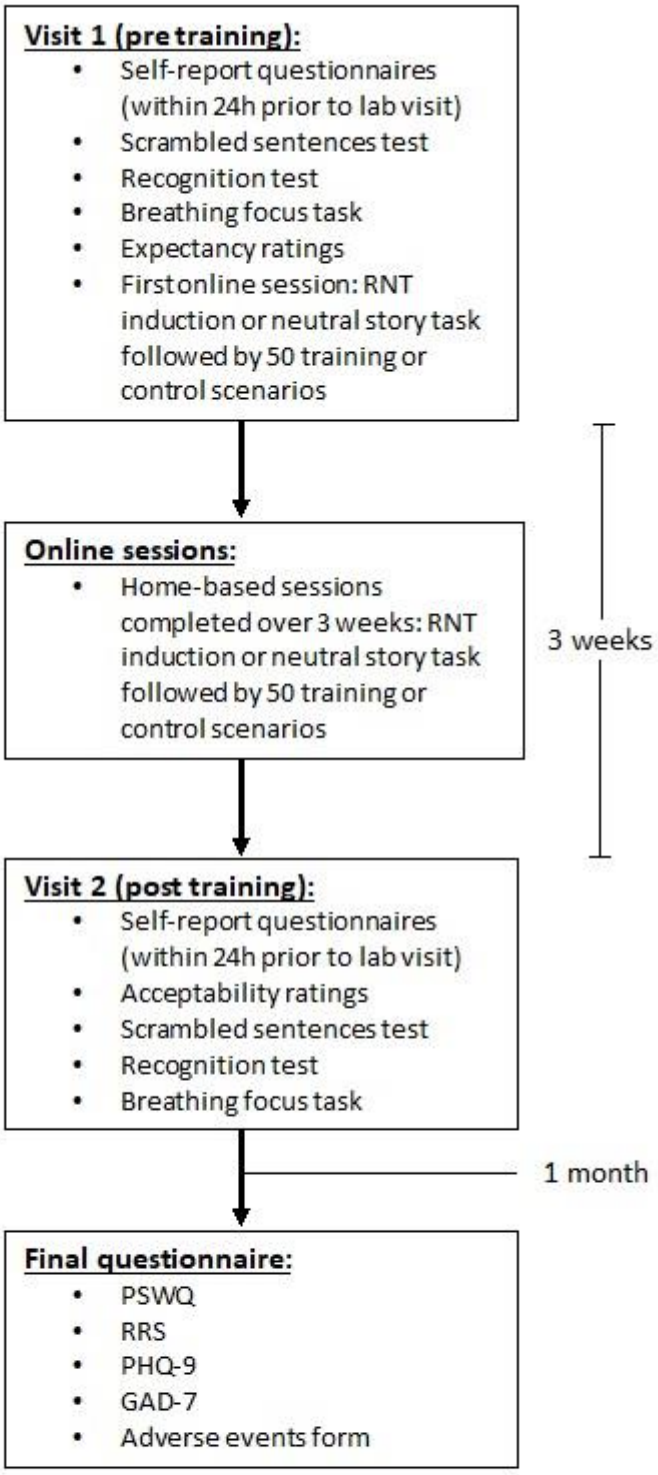

Figure 1. *Key elements of the procedure.*

Note. RNT = repetitive negative thinking; PSWQ = Penn State Worry Questionnaire; RRS = Ruminative Response Scale; PHQ-9 = Patient Health Questionnaire 9; GAD-7 = Generalized Anxiety Disorder 7-item scale.